

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

Point of Contact:
Jan Datsval
Librarian-Physical Sciences
CM1 1E01 Tel: 308-4498

STAFF USE ONLYSearcher: JanSearcher Phone #: 4458

Searcher Location: _____

Date Searcher Picked Up: 1/8Date Completed: 1/9

Searcher Prep & Review Time: _____

Clerical Prep Time: 10Online Time: 10**Type of Search**NA Sequence (#) ☒AA Sequence (#) ☒

Structure (#) _____

Bibliographic _____

Litigation _____

Fulltext _____

Patent Family _____

Other _____

Vendors and cost where applicable

STN _____

Dialog _____

Questel/Orbit _____

Dr.Link _____

Lexis/Nexis _____

Sequence Systems ☒

WWW/Internet _____

Other (specify) _____

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

protein - protein search, using sw model

on: January 8, 2002, 10:50:18 ; Search time 12.48 Seconds
(without alignments)
207.362 Million cell updates/sec

File: US-09-640-041-4
625
Sequence: 1 MPTDHEPCGSHKSFCLNG.....YDINLVETSTSAHSHQH 115

oring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

arched: 212252 seqs, 22503292 residues

tal number of hits satisfying chosen parameters: 212252

imum DB seq length: 0
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

tabase :
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	ID	Description
1	142	22.7	713	3	US-08-899-437-2
2	142	22.7	713	4	US-09-126-121-2
3	140	22.4	602	1	US-08-168-091A-2
4	140	22.4	602	1	US-08-428-926-5
5	140	22.4	602	1	US-08-428-927-5
6	140	22.4	602	1	US-08-428-928-5
7	140	22.4	602	1	US-08-339-517-5
8	140	22.4	1070	4	US-08-697-984-2
9	136	21.8	696	3	US-08-899-437-23
10	136	21.8	696	4	US-09-126-121-23
11	136	21.8	720	3	US-08-899-437-6
12	136	21.8	720	4	US-09-126-121-6
13	132.5	21.2	560	3	US-08-341-018-58
14	132.5	21.2	560	4	US-08-470-335-14
15	130.5	20.9	560	4	US-08-470-335-194
16	130.5	20.9	152	4	US-09-188-930-187
17	126.5	20.2	604	4	US-08-470-335-227
18	126.5	20.2	625	1	US-07-847-743B-26
19	126.5	20.2	625	1	US-08-456-201-26
20	126.5	20.2	625	2	US-08-456-241-26
21	126.5	20.2	625	3	PCT-US92-04295A-26
22	126.5	20.2	638	4	US-08-470-335-240
23	126.5	20.2	669	1	US-07-847-743B-8
24	126.5	20.2	669	1	US-07-847-743B-13
25	126.5	20.2	669	1	US-08-456-201-8
26	126.5	20.2	669	1	US-08-456-201-13
27	126.5	20.2	669	2	US-08-330-161-11

28	126.5	20.2	669	2	US-08-456-241-8	Sequence 8, Appl
29	126.5	20.2	669	2	US-08-456-241-13	Sequence 13, Appl
30	126.5	20.2	669	2	US-08-440-401-11	Sequence 11, Appl
31	126.5	20.2	669	2	US-08-419-878B-11	Sequence 11, Appl
32	126.5	20.2	669	5	PCT-US92-04295A-8	Sequence 8, Appl
33	126.5	20.2	669	5	PCT-US92-04295A-13	Sequence 13, Appl
34	126.5	20.2	821	4	US-08-470-335-228	Sequence 228, App
35	126.5	20.2	855	4	US-08-470-335-241	Sequence 241, App
36	126.5	20.2	868	4	US-08-470-335-229	Sequence 229, App
37	126.5	20.2	902	4	US-08-470-335-242	Sequence 242, App
38	124.5	19.9	121	1	US-08-549-757A-14	Sequence 14, Appl
39	124	19.8	420	1	US-07-847-743B-29	Sequence 29, Appl
40	124	19.8	420	1	US-08-456-201-29	Sequence 29, Appl
41	124	19.8	420	2	US-08-456-241-29	Sequence 29, Appl
42	124	19.8	420	5	PCT-US92-04295A-29	Sequence 29, Appl
43	124	19.8	601	4	US-08-470-335-233	Sequence 233, Appl
44	124	19.8	635	4	US-08-470-335-247	Sequence 247, App
45	124	19.8	637	1	US-07-847-743B-28	Sequence 28, Appl

ALIGNMENTS

```
RESULT 1
US-08-899-437-2
; Sequence 2, Application US/08899437
; Patent No. 6121415
;
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/899,437
; FILING DATE: 24-Jul-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Conley, Deirdre L.
; REGISTRATION NUMBER: 36,487
; REFERENCE/DOCKET NUMBER: P1084R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-2066
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 713 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
; FEATURE:
; NAME/KEY: Mouse NRG3 (mNRG3)/amino acid seq.
; LOCATION: 1-713
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
;
; US-08-899-437-2
;
; Query Match 22.7% Score 142; DB 3; Length 713;
; Best Local Similarity 26.6%; Pred. No. 6.1e-08;
; Matches 38; Conservative 20; Mismatches 47; Indels 38; Gaps 6;
;
; 3 TDHEPCGSHKSFCLNGLCYVITIPS--PFCRCVENYTGARCEVFLP----- 51
```

BASE COUNT 90 a 85 c 80 g 93 t

Query Match 100.0%; Score 348; DB 6; Length 348;
Best Local Similarity 100.0%; Pred. No. 4.3e-103; Indels 0; Gaps 0;
Matches 348; Conservative 0; Mismatches 0;

1 atgccacagatcacgaagaccctgtgtgtccagtcacaaagtcgtttgctgaatgg 60
|||||
1 atgccacagatcacgaagaccctgtgtgtccagtcacaaagtcgtttgctgaatgg 60
1 atgccacagatcacgaagaccctgtgtgtccagtcacaaagtcgtttgctgaatgg 60
61 ggccctgtgtatgtatgacttaccatctccacgcttggatgtgctgtaaaactat 120
|||||
61 ggccctgtgtatgtatgacttaccatctccacgcttggatgtgctgtaaaactat 120
121 aaagagagcctgtgtgaagaagttttccccaagctccagatccaactaaagtaac 180
|||||
121 aaagagagcctgtgtgaagaagttttccccaagctccagatccaactaaagtaac 180
121 acagagagcctgtgtgaagaagttttccccaagctccagatccaactaaagtaac 180
|||||
121 acagagagcctgtgtgaagaagttttccccaagctccagatccaactaaagtaac 180
181 ctgttgaagccttctgtgcatgtgcgtccagtaaacattatcatgtgagcctctac 240
|||||
181 ctgttgaagccttctgtgcatgtgcgtccagtaaacattatcatgtgagcctctac 240
241 ttcccttgcagaagaagccactttcagagaagccagttcagtcagatgatatcaacctg 300
|||||
241 ttcccttgcagaagaagccactttcagagaagccagttcagtcagatgatatcaacctg 300
301 gttagagagcagcagctacagctgtcccccacagtcattgaacaactga 348
|||||
301 gttagagagcagcagctacagctgtcccccacagtcattgaacaactga 348
301 gttagagagcagcagctacagctgtcccccacagtcattgaacaactga 348

RESULT 2
AX088170 348 bp DNA PAT 17-MAR-2001
LOCUS AX088170 Sequence 1 from Patent WO0114415.
ACCESSION AX088170
VERSION AX088170.1 GI:13397083
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE Kavanagh, W.M., Gen. H. and Lee, P.
AUTHORS EGF2 genes and gene products.
TITLE Patent: WO 0114415-A 1 01-MAR-2001;
JOURNAL CHIRON CORPORATION (US)
FEATURES
source 1. 348
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 82 a 102 c 87 g 77 t
ORIGIN

Query Match 72.4%; Score 252; DB 6; Length 348;
Best Local Similarity 82.8%; Pred. No. 1.2e-71;
Matches 288; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
1 atgccacagatcacgaagaccctgtgtgtccagtcacaaagtcgtttgctgaatgg 60
|||||
1 atgccacagatcacgaagaccctgtgtgtccagtcacaaagtcgtttgctgaatgg 60
1 atgccacagatcacgaagaccctgtgtgtccagtcacaaagtcgtttgctgaatgg 60
61 ggccctgtgtatgtatgacttaccatctccacgcttggatgtgctgtaaaactat 120
|||||
61 ggccctgtgtatgtatgacttaccatctccacgcttggatgtgctgtaaaactat 120
121 aaagagagcctgtgtgaagaagttttccccaagctccagatccaactaaagtaac 180
|||||
121 aaagagagcctgtgtgaagaagttttccccaagctccagatccaactaaagtaac 180
121 acagagagcctgtgtgaagaagttttccccaagctccagatccaactaaagtaac 180
|||||
121 acagagagcctgtgtgaagaagttttccccaagctccagatccaactaaagtaac 180

QY 181 ctgttgaagccttctgtgcatgtgcgtccagtaaacattatcatgtgagcctctac 240
|||||
DB 181 ctgttgaagccttctgtgcatgtgcgtccagtaaacattatcatgtgagcctctac 240
QY 241 ttcccttgcagaagaagccactttcagagaagccagttcagtcagatgatatcaacctg 300
|||||
DB 241 ttcccttgcagaagaagccactttcagagaagccagttcagtcagatgatatcaacctg 300
QY 301 gttagagagcagcagctacagctgtcccccacagtcattgaacaactga 348
|||||
DB 301 gttagagagcagcagctacagctgtcccccacagtcattgaacaactga 348

RESULT 3
AF083067 750 bp mRNA ROD 03-JUN-1999
LOCUS AF083067 Mus musculus neuregulin-4 short isoform (Nrg4) mRNA, complete cds.
DEFINITION AF083067.2 GI:4539680
ACCESSION AF083067.2 GI:4539680
VERSION AF083067.2 GI:4539680
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE Harari, D., Zohar, E., Romano, J., Shelly, M., Pierce, J.H.,
AUTHORS Harari, D., Zohar, E., Romano, J., Shelly, M., Pierce, J.H.,
Andrews, G.C. and Yarden, Y.
TITLE Neuregulin-4: a novel growth factor that acts through the ErbB-4
receptor tyrosine kinase
JOURNAL Oncogene 18 (17), 2681-2689 (1999)
MEDLINE 99276098
PUBMED 10348342
REFERENCE 2 (bases 1 to 750)
AUTHORS Harari, D. and Yarden, Y.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-1998) Biological Regulation, The Weizmann
Institute of Science, Rehovot 76100, Israel
COMMENT On Mar 30, 1999 this sequence version replaced gi:4530129.
FEATURES
source 1. 750
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
gene 1. 750
/gene="Nrg4"
64..411
/gene="Nrg4"
/note="ErbB-4 ligand"
/codon_start=1
/product="neuregulin-4 short isoform"
/protein_id="AAD21874.1"
/db_xref="GI:4530130"
/translation="MPTDHPDPCGRHRSFCINGICVIVIPSPFCINMYGAR
CEEVFLPSISPSNSLSAAFLVLAVALLTIALFLCKRKLORRASSVOCEISLVE
TNNRTFRSHREH"
CDS
64..411
/gene="Nrg4"
/note="ErbB-4 ligand"
BASE COUNT 209 a 191 c 164 g 186 t
ORIGIN

Query Match 72.4%; Score 252; DB 10; Length 750;
Best Local Similarity 82.8%; Pred. No. 1.4e-71;
Matches 288; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
1 atgccacagatcacgaagaccctgtgtgtccagtcacaaagtcgtttgctgaatgg 69
|||||
1 atgccacagatcacgaagaccctgtgtgtccagtcacaaagtcgtttgctgaatgg 69
1 atgccacagatcacgaagaccctgtgtgtccagtcacaaagtcgtttgctgaatgg 69
61 ggccctgtgtatgtatgacttaccatctccacgcttggatgtgctgtaaaactat 120
|||||
61 ggccctgtgtatgtatgacttaccatctccacgcttggatgtgctgtaaaactat 120
DB 124 ggcatgtgtatgtatgacttaccatctccacgcttggatgtgctgtaaaactat 183
|||||
QY 121 aaagagagcctgtgtgaagaagttttccccaagctccagatccaactaaagtaac 180
|||||

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

protein - protein search, using sw model

on: January 8, 2002, 10:49:43; Search time 23.58 Seconds
(without alignments)
361.257 Million cell updates/sec

le: US-09-640-041-4

fect score: 625
1 MPTDHEPCGSPSHKSFCLNG.....YDINLVETSTSAHSHSECH 115

ing table: BLOSUM62
Gapop 10.0, Gapext 0.5

ched: 522463 segs, 74073290 residues

cal number of hits satisfying chosen parameters: 522463

imum DB seq length: 0

imum DB seq length: 2000000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase :
1: /SID88/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SID88/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SID88/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SID88/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SID88/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SID88/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SID88/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SID88/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SID88/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SID88/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SID88/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SID88/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SID88/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SID88/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SID88/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SID88/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SID88/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SID88/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SID88/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SID88/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SID88/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SID88/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Sult No.	Score	Query Match	Length	DB ID	Description
1	625	100.0	115	22	AA97697
2	489	78.2	115	22	AA97698
3	142	22.7	713	20	AA97617
4	140	22.4	602	15	AA97617
5	140	22.4	602	15	AA97617
6	140	22.4	602	15	AA97617
7	140	22.4	602	15	AA97617
8	136	21.8	1070	22	AA97619
9	136	21.8	720	20	AA97618
10	132.5	21.2	560	15	AA97618
11	132.5	21.2	560	15	AA97618

12	132.5	21.2	560	16	AA976245
13	132.5	21.2	560	17	AA976245
14	132.5	21.2	560	17	AA976245
15	132.5	21.2	560	17	AA976245
16	132.5	21.2	560	20	AA976245
17	130.5	20.9	152	21	AA976245
18	130.5	20.9	152	22	AA976245
19	130	20.8	675	19	AA976245
20	130	20.8	675	19	AA976245
21	129	20.6	675	19	AA976245
22	129	20.6	675	19	AA976245
23	128	20.5	637	13	AA976245
24	127	20.3	675	20	AA976245
25	127	20.3	675	21	AA976245
26	126.5	20.2	462	16	AA976245
27	126.5	20.2	462	16	AA976245
28	126.5	20.2	625	20	AA976245
29	126.5	20.2	669	13	AA976245
30	126.5	20.2	669	13	AA976245
31	126	20.2	675	21	AA976245
32	124.5	19.9	121	16	AA976245
33	124	19.8	276	16	AA976245
34	124	19.8	420	13	AA976245
35	124	19.8	420	20	AA976245
36	124	19.8	420	21	AA976245
37	124	19.8	637	20	AA976245
38	124	19.8	637	21	AA976245
39	124	19.8	675	19	AA976245
40	124	19.8	675	21	AA976245
41	124	19.8	675	21	AA976245
42	124	19.8	675	21	AA976245
43	124	19.8	675	21	AA976245
44	123.5	19.8	298	22	AA976245
45	123.5	19.8	422	14	AA976245

ALIGNMENTS

RESULT 1
ID AA97697 standard; Protein: 115 AA.
AC AA97697;
DT 19-JUN-2001 (first entry)
DE Human EGFH2 protein sequence.
XX EGFH2; epidermal growth factor H2; EGFH2 protein modulated disorder;
KW cancer; peripheral neuropathy; amyotrophic lateral sclerosis; therapy;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW ischaemic stroke; brain injury; acute spinal cord injury; infection;
KW nervous system injury; multiple sclerosis; dementia; epilepsy;
KW peripheral nerve injury; acoustic trauma; human.
XX Homo sapiens.
OS
PN WO200114415-A2.
PD 01-MAR-2001.
PF 15-AUG-2000; 2000WO-US22326.
PR 20-AUG-1999; 99US-0149986.
PA (CHIR) CHIRON CORP.
XX Kavanagh WM, Cen H, Lee P;
XX WPI: 2001-211304/21.
DR N-PSDB; AA971294.
XX

Bovine glial cell
Human neuroglial G
Glial growth facto
BPP4 glial growth
Amino acid sequenc
Murine TGF-alpha h
Skin cell protein,
Amino acid sequenc
Amino acid sequenc
Human heregulin HR
Human heregulin-be
Human heregulin-11
Human heregulin HR
Human NDF-alpha2b
Human PRONDF-alpha
Human heregulin-11
Human heregulin-al
Human heregulin HR
Human heregulin HR
HCT-15 contg. HTG-
Human PRONDF-beta2
Human heregulin-be
Human heregulin-beta-2-1
Human heregulin HR
Human heregulin-beta-2.
Human heregulin HR
Amino acid sequenc
Human heregulin HR
Human heregulin HR
Human heregulin HR
Human heregulin HR
Amino acid sequenc
Rat NRSE. Rattus

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

protein - protein search, using sw model

on: January 8, 2002, 10:58:38 ; Search time 22.24 Seconds

(without alignments)
756.354 Million cell updates/sec

File: us-09-640-041-4

quence score: 1 MPTDHEPCGSHKSFCLNG.....YDINLVETSTSAHSHSQH 115

oring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 473505 segs, 146272329 residues

11 number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Minimum DB seq length: 2000000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP_ARCHA:.*
2: SP_BACTERIA:.*
3: SP_FUNGI:.*
4: SP_HUMAN:.*
5: SP_INVERTEBRATE:.*
6: SP_MAMMAL:.*
7: SP_MNC:.*
8: SP_ORGANELLAE:.*
9: SP_PHAGE:.*
10: SP_PLANT:.*
11: SP_RODENT:.*
12: SP_VIRUS:.*
13: SP_VERTEBRATE:.*
14: SP_UNCLASSIFIED:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Sult No.	% Query Match	Length	DB ID	Description
1	130.5	20.9	461 11 035947	035947 mesocricetu
2	128.5	20.6	1214 6 09BDQ0	09BDQ0 sus scrofa
3	123.5	19.8	298 11 09ES49	09ES49 rattus norv
4	123.5	19.8	695 11 09ESB0	09ESB0 rattus norv
5	122.5	19.6	162 11 061521	061521 mus musculu
6	118.5	19.0	162 11 092055	092055 rattus norv
7	117.5	18.8	127 11 0909F3	0909F3 ectromelia
8	116	18.6	700 11 09ESB1	09ESB1 rattus norv
9	116	18.6	782 11 09ES45	09ES45 rattus norv
10	115	18.4	140 12 089756	089756 variola vir
11	113.5	18.2	127 12 041506	041506 bean 58058
12	113.5	18.2	140 12 057166	057166 vaccinia vi
13	113.5	18.2	140 12 09UR4	09UR4 vaccinia vi
14	112	17.9	140 12 089066	089066 variola vir
15	111.5	17.8	140 12 086607	086607 vaccinia vi
16	109.5	17.5	138 12 087605	087605 compox viru
17	109.5	17.5	241 6 007112	007112 bos taurus
18	108.5	17.4	125 12 041504	041504 compox viru
19	107.5	17.2	111 11 09ES48	09ES48 rattus norv

20	107.5	17.2	136 11 09ES47	09ES47 rattus norv
21	107.5	17.2	236 11 09ES46	09ES46 rattus norv
22	107.5	17.2	317 11 09ES43	09ES43 rattus norv
23	107.5	17.2	323 11 09ES42	09ES42 rattus norv
24	107.5	17.2	342 11 09ES41	09ES41 rattus norv
25	106.5	17.0	642 13 079941	079941 xenopus lae
26	105.5	16.9	295 5 09VL2	09VL2 drosophila
27	105.5	16.9	4599 5 09NZR2	09NZR2 homo sapien
28	105	16.8	212 13 09W7C5	09W7C5 gallus gall
29	103.5	16.6	4601 5 09V383	09V383 drosophila
30	103	16.5	247 6 09BDH4	09BDH4 sus scrofa
31	101.5	16.2	4599 11 09J118	09J118 mus musculu
32	100	16.0	707 5 09VR03	09VR03 drosophila
33	99.5	15.9	551 13 013003	013003 gallus gall
34	98.5	15.8	159 4 09BS56	09BS56 homo sapien
35	98.5	15.8	161 4 09UKM1	09UKM1 homo sapien
36	98.5	15.8	162 4 09UKM2	09UKM2 homo sapien
37	98.5	15.8	163 4 09UKM3	09UKM3 homo sapien
38	96	15.4	188 11 09ROC7	09ROC7 ciliate
39	96	15.4	1643 6 077611	077611 bos taurus
40	95	15.2	105 11 099NM7	099NM7 mesocricetu
41	95	15.2	746 4 09H1R1	09H1R1 homo sapien
42	94	15.0	89 6 028867	028867 equus cabal
43	94	15.0	656 6 077612	077612 bos taurus
44	94	15.0	1372 5 091526	091526 caenorhabd
45	94	15.0	2394 6 077610	077610 bos taurus

ALIGNMENTS

RESULT 1
ID 035947 AC 035947 PRELIMINARY: PRT: 461 AA.
DT 01-JAN-1998 (TRENBLER, 05, Created)
DT 01-JAN-1998 (TRENBLER, 05, Last sequence update)
DT 01-JUN-2001 (TRENBLER, 17, Last annotation update)
DE PRO-NEUREGULIN-1, ISOFORM ALPHA 2b PRECURSOR.
GN NRGI OR NDE.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA2B), AND SEQUENCE OF 64-81.
RC TISSUE=EMBRYO;
RX MEDLINE=98196996; PubMed=9537646;
RA Velasco J.A., Feljoo E., Avila M.A., Notario V.;
RT "Secretion of neu differentiation factor-like polypeptides by cph-
transformed fibroblasts: cloning and characterization of Syrian
hamster neuregulin cDNAs";
RT Mol. Carcinog. 21:156-163(1998).
RL
CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORCEPTORS,
RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
ACTIVATION OF THE ERBB RECEPTORS. MAY PLAY AN IMPORTANT ROLE IN
PROVIDING GROWTH ADVANTAGE IN NEOPLASTIC CELLS.
CC -1- SUBUNIT: THE CYTOPLASMIC DOMAIN INTERACTS WITH THE LIM DOMAIN
REGION OF LIMK1 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: EXISTS AS TYPE I MEMBRANE PROTEIN AND AS A
PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-
BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: DIFFERENT ISOFORMS ARE PRODUCED BY
ALTERNATIVE SPLICING. THE SEQUENCE SHOWN IS THAT OF ISOFORM
ALPHA2B/CLONE PM3.
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHER LEVEL AFTER NEOPLASMIC
TRANSFORMATION OF CELLS.
CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
DIMERIZATION (BY SIMILARITY).

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

protein - protein search, using sw model

on: January 8, 2002, 10:58:58 ; Search time 11.71 Seconds

(without alignments)
360.073 Million cell updates/sec

file: US-09-640-041-4

reflect score: 625

sequence: 1 MPTDHEPCGSHKSFCLNG.....YDINLVETSTAHSHSDH 115

oring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

arched: 100059 seqs, 36664827 residues

tal number of hits satisfying chosen parameters: 100059

imum DB seq length: 0

imum DB seq length: 200000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

tabase : SwissProt_39:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	Query Match Length	ID	Description
1	489	78.2	115	NRG4_MOUSE
2	142	22.7	713	Q95181 mus musculus
3	140	22.4	602	NRG1_CHICK
4	136	21.8	720	NRG3_HUMAN
5	126.5	20.2	639	NRG1_HUMAN
6	124.5	19.9	169	NRG1_MOUSE
7	119.5	19.1	1217	NRG2_MOUSE
8	119	19.0	756	EGF_MOUSE
9	116.5	18.6	1133	EGF_MOUSE
10	115	18.4	140	EGF_MOUSE
11	113	18.1	53	EGF_MOUSE
12	113	18.1	4655	LRP2_HUMAN
13	112	17.9	858	NRG2_HUMAN
14	112	17.9	868	NRG2_MOUSE
15	109.5	17.5	296	SMDF_HUMAN
16	109	17.4	1207	EGF_MOUSE
17	108.5	17.4	140	EGF_MOUSE
18	108.5	17.2	677	NRG1_MOUSE
19	107.5	17.2	662	NRG1_MOUSE
20	105.5	16.9	294	GRK_MOUSE
21	103.5	16.6	142	GRK_MOUSE
22	102.5	16.4	85	GRK_MOUSE
23	101.5	16.2	208	GRK_MOUSE
24	101	16.2	4543	LRP1_MOUSE
25	100.5	16.1	80	VEIN_MOUSE
26	100	16.0	622	VEIN_MOUSE
27	99.5	15.9	133	TGFA_MOUSE
28	99.5	15.9	160	TGFA_MOUSE
29	99.5	15.9	208	TGFA_MOUSE
30	99.5	15.9	208	TGFA_MOUSE
31	98.5	15.8	160	TGFA_MOUSE
32	98.5	15.8	208	TGFA_MOUSE
33	97.5	15.6	121	TGFA_MOUSE

34	96.5	15.4	208	1	HRGF_RAT	006175	rattus norv
35	96	15.4	178	1	BRC_HUMAN	P35070	homo sapien
36	95.5	15.3	178	1	BRC_BOVIN	Q91C5	bos taurus
37	93	14.9	402	1	LAG2_CAEL	P45442	caenorhabd
38	93	14.9	2139	1	CRB_MOUSE	P10040	drosophila
39	92	14.7	746	1	MEPA_HUMAN	Q16819	homo sapien
40	91.5	14.6	159	1	TGFA_MOUSE	P48030	mus musculus
41	91.5	14.6	159	1	TGFA_MOUSE	P01134	rattus norv
42	91	14.6	177	1	BRC_MOUSE	Q05928	mus musculus
43	90.5	14.5	248	1	SDGF_MOUSE	P31955	mus musculus
44	90.5	14.5	252	1	AMPR_HUMAN	P15514	homo sapien
45	90	14.4	862	1	PGCV_MACNE	Q28858	macaca neme

ALIGNMENTS

RESULT 1
NRG4_MOUSE STANDARD: PRT: 115 AA.
ID NRG4_MOUSE
AC Q9WTF4;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PRO-NEUREGULIN-4, SHORT ISOFORM (PRO-NRG4) [CONTAINS: NEUREGULIN-4 (NRG-4)]
DE NRG4.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
OX NCBT_TaxID=10090;
RX MEDLINE=99276098; PubMed=10348342;
RA Harari D., Tzahar E., Romano J., Shelly M., Pierce J.H., Andrews G.C.,
RA Yarden Y.;
RT "Neuregulin-4: a novel growth factor that acts through the ErbB-4
RT receptor tyrosine kinase."
RL Oncogene 18:2681-2689(1999).
CC -1- FUNCTION: LOW AFFINITY LIGAND FOR THE ERBB4 TYROSINE KINASE
CC RECEPTOR. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
CC ACTIVATION OF THE ERBB RECEPTORS. DOES NOT BIND TO THE ERBB1,
CC ERBB2 AND ERBB3 RECEPTORS.
CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS
CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS MAY BE PRODUCED BY
CC ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PANCREAS; WEAKLY EXPRESSED
CC IN MUSCLE.
CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC DIMERIZATION (BY SIMILARITY).
CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC DOMAIN (BY SIMILARITY).
CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC FORM (BY SIMILARITY).
CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).


```

ADDRESS: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatlin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,437
FILING DATE: 24-Jul-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Delidre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: PI084RL
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2538 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear

FEATURE:
NAME/KEY: mouse NR3 nucleic acid
LOCATION: 1-2538
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-08-899-437-1

```

Query Match	8.6%	Score 29.8	DB 3	length 2538
Best Local Similarity	55.0%	Pred. No.	1.3	
Matches	83	Conservative	0	Mismatches 62
				Indels 6
				Gaps 2

[illegible]

RESULT 3
US-09-126-121-1

Sequence 1 Application US/09126121
Patent No. 6252051
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
TITLE OF INVENTION: Ligands and Uses Therefor
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/126,121
FILING DATE: 30-Jul-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Deidre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2538 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
FEATURE:
NAME/KEY: name NRG3 nucleic acid
LOCATION: 1-2538
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-09-126-121-1

```

Query Match	8.6%	Score 29.8	DB 4	length 2538
Best Local Similarity	55.0%	Pred. 1.3		
Matches 83	Conservative	0	Mismatches 62	Indels 6
				Gaps 1

QY 4 ccaacagatcagaagaagccctgctggtcccaagtcacaagtcgttttgcctgaatgagggg 63
Db 1141 CGATCTGAGCACTTCAACCCCTGTCCAGACAAGAGACCTGGCTATGTCTCAATGATGT 1200
QY 64 ctttgtatgtgtacctactatctccagccattt-----tgtagtggtgtgaaac 117
Db 1201 GAATGCTTGTGTGATTGAGACCCCTGAGAAGGATCCCATTAACCACTGTGGTGCAAGGAAGC 1260
QY 118 tatcaggaagctcgtttgtgaagaggttttc 148
Db 1261 TACCAGAAGAGTCCGTTGTGTGATCAATTTCTGC 1291

RESULT 4
US-08-482-918-47

Sequence 47, Application US/0842918
Patent No. 6207417
GENERAL INFORMATION:
APPLICANT: Zsebo, Kristzina M.
APPLICANT: Bosselman, Robert A.
APPLICANT: Suggs, Sidney V.
APPLICANT: Martin, Francis H.
TITLE OF INVENTION: Stem Cell Factor
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borum
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago

STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,918
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

protein - protein search, using sw model

on: January 8, 2002, 10:56:28 ; Search time 14.34 Seconds

(without alignments)
610.884 Million cell updates/sec

file: US-09-640-041-4

effect score: 1 MPDHEPCGSPSHKSPCLNG.....YDINLVETSTSAHSHQEH 115

quence: 1 MPDHEPCGSPSHKSPCLNG.....YDINLVETSTSAHSHQEH 115

oring table: BROSUM62
Gap: 10.0 , Gapext 0.5

arched: 219241 seqs, 76174552 residues

total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Minimum DB seq length: 2000000000

st-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

- 1: PIR68:*
- 2: PIR1:*
- 3: PIR3:*
- 4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Sult No.	Score	Query Match	Length	DB ID	Description
1	142	22.7	713	2	neuregulin-3 [Impo
2	140	22.4	602	2	acetylcholine rece
3	126.5	20.2	462	2	neu differentiation
4	126.5	20.2	640	2	neuregulin precurs
5	124	19.8	637	2	neuregulin precurs
6	123.5	19.8	639	2	neu differentiation
7	123	19.7	645	2	neuregulin, splice
8	122.5	19.6	162	2	epiregulin, precus
9	119.5	19.1	1217	1	epidermal growth f
10	117	18.7	636	2	neu differentiation
11	116.5	18.6	125	2	neuregulin isoform
12	116.5	18.6	125	2	neu differentiation
13	116.5	18.6	125	2	epidermal growth f
14	115	18.4	140	2	hypothetical prote
15	115	18.4	140	2	hypothetical prote
16	113.5	18.2	140	2	epidermal growth f
17	113	18.1	53	2	growth factor - va
18	112	17.9	140	2	B3R protein - vari
19	112	17.9	621	2	low density lipopr
20	112	17.9	850	2	erbB kinase activa
21	112	17.9	860	2	erbB kinase activa
22	112	17.9	868	2	erbB kinase activa
23	109.5	17.5	175	2	neu differentiation
24	109.5	17.5	241	2	glial growth facto
25	109.5	17.5	241	2	neuregulin precurs
26	109.5	17.5	296	2	sensory/motor neur
27	109.5	17.5	422	2	glial growth facto
28	109	17.4	1207	1	epidermal growth f
29	108.5	17.4	140	1	growth factor - va

30	107.5	17.2	230	2	A56210	neu differentiation
31	107.5	17.2	662	2	I61722	neu differentiation
32	105.5	16.9	294	2	A4844	TGF alpha-like pro
33	103.5	16.6	142	2	WMV23C	growth factor - va
34	102.5	16.4	85	1	EGV2M1	growth factor - my
35	101	16.2	4543	1	A51102	alpha-2-macroglobu
36	100.5	16.1	80	1	EGV2SF	growth factor - ra
37	99.5	15.9	82	2	A61040	transforming growt
38	99.5	15.9	160	2	S39795	transforming growt
39	99.5	15.9	208	1	A38432	heparin-binding EG
40	99.5	15.9	208	1	A41914	heparin-binding EG
41	98.5	15.8	160	1	WPHU1	transforming growt
42	98.5	15.8	208	1	JC1410	transforming growt
43	97.5	15.6	121	2	I58134	transforming growt
44	96.5	15.4	208	1	JC1409	heparin-binding EG
45	96	15.4	178	2	JC1467	betacellulin precu

ALIGNMENTS

RESULT 1
T44447
neuregulin-3 [Imported] - mouse
C:Species: Mus musculus (house mouse)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
R:Zhang, D.; Sliwkowski, M.X.; Mark, M.; Frantz, G.; Akita, R.; Sun, Y.; Hillan, K.;
Proc. Natl. Acad. Sci. U.S.A. 94, 9562-9567, 1997
A:Title: Neuregulin-3 (NRG3): A novel neural tissue-enriched protein that binds and s
A:Reference number: 222773; MUID:97420720
A:Accession: T44447
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-713 <ZHA>
A:Cross-references: EMBL:AF010130; NID:g2429163; PIDN:AAB70914.1; PID:g2429164
A:Gene: NRG3

Query Match 22.7%; Score 142; DB 2; Length 713;
Best Local Similarity 26.6%; Pred. No. 6.1e-07;
Matches 38; Conservative 20; Mismatches 47; Indels 36; Gaps 6;
QY 3 TDHEPCGSPSHKSPCLNGSLCYVPIIPS--PFCRCVENYTGARCEVFLP----- 51
DB 286 SEHFPCKRCKDLAYCLNDECFYETLTGSHKRCCKEGVGRCDQ-FLPKTDSILSDP 344
QY 52 -----GSSIQTKSNLE-----AFVALAVLVTLIGATYFLCRKGHPORASSVQ----- 95
DB 345 TDHLGTFEMSESDVYORQVLSICJIFGIVGMFCACAFYFSKR---QAKOIOEHLKE 400
QY 96 -----YDINLVETSTSAHSH 112
DB 401 SONGKNYSLKASTKSESLSMKSH 423
RESULT 2
A45769
acetylcholine receptor synthesis stimulator ARIA-1 precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 21-Jul-2000
R:Fallis, D.L.; Rosen, K.M.; Cortas, G.; Lane, W.S.; Fischbach, G.D.
Cell 72, 801-815, 1993
A:Title: ARIA, a protein that stimulates acetylcholine receptor synthesis, is a membe
A:Reference number: A45769; MUID:93201602
A:Accession: A45769
A:Status: Preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-602 <FAI>
A:Cross-references: GB:J11264; NID:g212603; PIDN:AAA49037.1; PID:g212604
A:Experimental source: Brain

A:Note: sequence extracted from NCBI backbone (NCBIN:127787, NCBIPI:127788)
C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 22.4%; Score 140; DB 2; Length 602;
Best Local Similarity 27.1%; Pred. No. 8.3e-07;

Matches 35; Conservative 21; Mismatches 51; Indels 22; Gaps 4;

3 TDHEPCGSHKSFCLNGLCYIPIPTSP---FCRCVENYTGARCEV-----48

DB 135 TSHLTCKDIQKAFVCGECYMKDLPNPRILCRCPNEFTGDRCONVMASFYKHLGI 194

49 -FLPGSIQKSNLFEAFVALAVLVLLIGAFYFLC-----RKGFQRASSVOYDINLV 101

DB 195 EFMAMELYOKRVLITIGICIALLVGIMCVAV-CYTKQKRRKLDRLQSLRSENNV 253

QY 102 ETSSTSAHH 110

DB 254 MNMANGPHH 262

RESULT 3

new differentiation factor - human

C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 05-Nov-1999

A:Accession: I38404

R:Men, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.;

Mol. Cell. Biol. 14, 1909-1919, 1994

A:Title: Structural and functional aspects of the multiplicity of Neu differentiation fa

A:Reference number: A56210; MUID:94158863

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1462 <RES>

A:Cross-references: EMBL:U02326; NID:g408402; PIDN:AAA19951.1; PID:g408403

Query Match 20.2%; Score 126.5; DB 2; Length 462;
Best Local Similarity 25.0%; Pred. No. 1.6e-05;

Matches 32; Conservative 25; Mismatches 46; Indels 25; Gaps 6;

3 TDHEPCGSHKSFCLNGLCYIPIPTSP---FCRCVENYTGARCEVFLPGSSITQTS 59

DB 176 TSHLVCAEKKEKFCVNGECFVWKDLSNPSRYLCKQPGFTGARCTE-NVP-MKYQNOE 233

QY 60 NLEFAFVALAVLT-----LIIGAFYFLC-----RKGFQRASSVOYDINLV 102

DB 234 KAEELYOKRVLITIGICIALLVG---IMCVAVCYTKQKRRKLDRLQSLRSENNM 290

QY 103 TSSTSAHH 110

DB 291 NIANGPHH 298

RESULT 4

herregulin precursor, splice form alpha - human

N:Alternate names: breast cancer cell differentiation factor p45; Neu differentiation fa

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 11-Jan-2000

C:Accession: A43273; A48498; A38155

R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yansu

Science 256, 1205-1210, 1992

A:Title: Identification of herregulin, a specific activator of p185(erbB2).

A:Reference number: A43273; MUID:92271253

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-640 <HOLD>

A:Experimental source: breast tumor cell line, MDA-MB-231, ATCC HTB 26

A:Note: sequence extracted from NCBI backbone (NCBIPI:103250)

R:Cilouscu, J.M.; Plovman, G.D.; Carlton, G.W.; Green, J.M.; Shoyab, M.

J. Biol. Chem. 268, 18407-18410, 1993

A:Title: Characterization of a breast cancer cell differentiation factor that speci

A:Reference number: A48498; MUID:9336731

A:Accession: A48498

A:Molecule type: protein

A:Residues: 20-21, 'X', 23-24, 'XX', 27-28 <CUI>

R:Peliss, E.; Bacus, S.S.; Koski, R.A.; Lu, H.S.; Wen, D.; Ogden, S.G.; Levy, R.B.; J

Cell 69, 205-216, 1992

A:Title: Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein that it

A:Reference number: A38155; MUID:92208945

A:Accession: A38155

A:Molecule type: protein

A:Residues: 'X', 15-16, 'X', 18-20, 'RG', 23-24, 'GP', 27, 'E', 29, 'XP', 32-36 <PEL>

A:Note: sequence extracted from NCBI backbone (NCBIPI:91347)

C:Genetics:

A:Gene: GDB:HGL

A:Cross-references: GDB:132656; OMIM:142445

A:Map position: 8p22-8p11

C:Superfamily: unassigned EGF-related proteins; EGF homology

C:Keywords: alternative splicing; glycoprotein

F:182-221/Domain: EGF homology <EGF>

Query Match 20.2%; Score 126.5; DB 2; Length 640;
Best Local Similarity 25.0%; Pred. No. 2.2e-05;

Matches 32; Conservative 25; Mismatches 46; Indels 25; Gaps 6;

3 TDHEPCGSHKSFCLNGLCYIPIPTSP---FCRCVENYTGARCEVFLPGSSITQTS 59

DB 176 TSHLVCAEKKEKFCVNGECFVWKDLSNPSRYLCKQPGFTGARCTE-NVP-MKYQNOE 233

QY 60 NLEFAFVALAVLT-----LIIGAFYFLC-----RKGFQRASSVOYDINLV 102

DB 234 KAEELYOKRVLITIGICIALLVG---IMCVAVCYTKQKRRKLDRLQSLRSENNM 290

QY 103 TSSTSAHH 110

DB 291 NIANGPHH 298

RESULT 5

herregulin precursor, splice form beta-2 - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Nov-2000

C:Accession: C43273; I38404

R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Y

Science 256, 1205-1210, 1992

A:Title: Identification of herregulin, a specific activator of p185(erbB2).

A:Reference number: A43273; MUID:92271253

A:Accession: C43273

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA

A:Residues: 1-637 <HOLD>

A:Experimental source: breast tumor cell line, MDA-MB-231, ATCC HTB 26

A:Note: sequence extracted from NCBI backbone (NCBIPI:103250)

R:Cilouscu, J.M.; Plovman, G.D.; Carlton, G.W.; Green, J.M.; Shoyab, M.

J. Biol. Chem. 268, 18407-18410, 1993

A:Title: Characterization of a breast cancer cell differentiation factor that speci

A:Reference number: A48498; MUID:9336731

A:Accession: A48498

A:Molecule type: protein

A:Residues: 20-21, 'X', 23-24, 'XX', 27-28 <CUI>

R:Peliss, E.; Bacus, S.S.; Koski, R.A.; Lu, H.S.; Wen, D.; Ogden, S.G.; Levy, R.B.; J

Cell 69, 205-216, 1992

A:Title: Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein that it

A:Reference number: A38155; MUID:92208945

A:Accession: A38155

A:Molecule type: protein

A:Residues: 'X', 15-16, 'X', 18-20, 'RG', 23-24, 'GP', 27, 'E', 29, 'XP', 32-36 <PEL>

A:Note: sequence extracted from NCBI backbone (NCBIPI:91347)

C:Genetics:

A:Gene: GDB:HGL

A:Cross-references: GDB:132656; OMIM:142445

A:Map position: 8p22-8p11

C:Superfamily: unassigned EGF-related proteins; EGF homology

C:Keywords: alternative splicing; glycoprotein

F:182-221/Domain: EGF homology <EGF>

Query Match

19.8%; Score 124; DB 2; Length 637;

NUMBER OF SEQUENCES: 23

```

-----
EMBL: AF083067; AAD21874.1;
InterPro: IPR0000561; EGF-like.
InterPro: IPR001336; EGF_1.
Pfam: PF00008; EGF, 1.
PRINTS: PR00009; EGF1GF.
SMART: SM00181; EGF, 1.
PROSITE: PS00022; EGF_1; 1.
PROSITE: PS01186; EGF_2; FALSE NEG.
Growth factor; EGF-like domain; Glycoprotein; Transmembrane;
Multigene family; Alternative splicing.
CHAIN 1 115
T T CHAIN 1 61
T T CHAIN 1 62
T T TRANSMEM 63 83
T T TRANSMEM 84 115
T T DOMAIN 5 46
T T DOMAIN 9 23
T T DISULFID 17 34
T T DISULFID 36 45
T T CARBOHYD 39 60
T T CARBOHYD 60 60
SEQUENCE 115 AA; 12743 MW; 989A1E376F857B49 CRC64;

Query Match 78.2%; Score 489; DB 1; Length 115;
Best Local Similarity 75.7%; Pred. No. 2.8e-45;
Matches 87; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

1 MPTDHEPCGPHSKSFCNLGCLCYVPTIPSPRCVENVYTGARCEVFLPSSSIOTKSN 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MPTDHPGCPGPHRSFCLNGCLCYVPTIPSPRCIEVYTGARCEVFLPSSSIPTSEN 60

61 LEFAFALAVLVTLIGAFYFLCRKGHPORASSVOYDINLVETSSSAHSHSQH 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 LSAFVAVLAVLTITLALCFCLCRKGLHORASSVOCEISLVETNNRTRSHREH 115

RESULT 2
NRG3_MOUSE STANDARD; PRT; 713 AA.
AC 035181;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PRO-NEUREGULIN-3 PRECURSOR (PRO-NRG3) [CONTAINS: NEUREGULIN-3 (NRG-
3)].
DE NRG3.
DE Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RT TISSUE-Brain;
RX MEDLINE=97420720; PubMed=9275162;
RX Zhang D., Sliwkowski M.X., Mark M., Fiantz G., Akita R., Sun Y.,
RX Hillan K., Crowley C., Brush J., Godowski P.J.;
RT "neuregulin-3 (NRG3): a novel neural tissue-enriched protein that
RT binds and activates ErbB4";
RT Proc. Natl. Acad. Sci. U.S.A. 94:9562-9567(1997)
CC -1- FUNCTION: DIRECT LIGAND FOR THE ERBB4 TYROSINE KINASE RECEPTOR.
CC BINDING RESULTS IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
CC ACTIVATION OF THE RECEPTOR. DOES NOT BIND TO THE EGF RECEPTOR,
CC ERBB2 OR ERBB3 RECEPTORS.
CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS
CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SYMPATHETIC, MOTOR, AND SENSORY
CC NEURONS.
CC -1- DEVELOPMENTAL STAGE: DETECTED AS EARLY AS E11. IN E13 EMBRYOS,
CC DETECTED MAINLY IN THE NERVOUS SYSTEM. IN E16 EMBRYOS, DETECTED IN
CC THE BRAIN, SPINAL CORD, TRIGEMINAL, VESTIBULAR-COCHLEAR, AND

```

```

CC SPINAL GANGLIA. IN ADULTS, EXPRESSED IN SPINAL CORD, AND NUMEROUS
CC BRAIN REGIONS.
CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC DIMERIZATION (BY SIMILARITY).
CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC DOMAIN (BY SIMILARITY).
CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC FORM (BY SIMILARITY).
CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF010130; AAB70914.1;
CC MGD: MGI:1097165; Nrg3.
CC DR InterPro: IPR0000561; EGF-like.
CC DR InterPro: IPR002154; Neuregulin.
CC DR Pfam: PF00008; EGF, 1.
CC DR Pfam: PF02158; Neuregulin; 1.
CC DR SMART: SM00181; EGF, 1.
CC DR PROSITE: PS00022; EGF_1; 1.
CC DR PROSITE: PS01186; EGF_2; 1.
CC DR Growth factor; EGF-like domain; Transmembrane; Multigene family.
CC KW CHAIN 1 713
CC FT CHAIN 1 361
CC FT DOMAIN 1 362
CC FT TRANSMEM 363 383
CC FT TRANSMEM 384 713
CC FT DOMAIN 105 287
CC FT DOMAIN 288 331
CC FT DOMAIN 13 21
CC FT DOMAIN 26 34
CC FT DOMAIN 127 135
CC FT DOMAIN 250 253
CC FT DOMAIN 254 263
CC FT DOMAIN 264 267
CC FT DISULFID 292 306
CC FT DISULFID 300 319
CC FT DISULFID 321 330
CC FT SEQUENCE 713 AA; 77369 MW; 9F7D1D5E7FC8DCE0 CRC64;

Query Match 22.7%; Score 142; DB 1; Length 713;
Best Local Similarity 26.6%; Pred. No. 1.4e-07;
Matches 38; Conservative 20; Mismatches 47; Indels 38; Gaps 6;

OY 3 TDHEPCGPHSKSFCNLGCLCYVPTIPSPRCVENVYTGARCEVFLP----- 51
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 286 SEHFKPCRDRLAVCLNDGCFVETLTGSHKRCRKGVGVCDO-FLPKTDSILSDP 344

OY 52 -----GSSIQTKSNLFE-----AFVALAVLVTLIGAFYFLCRKGHPORASSVO----- 95
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 345 TDHLGTFMESEDYQOVLSISCIITGIVAGFCAFYKRSK-----QAKQDEHLKE 400

OY 96 -----YDINLVETSSSAHSH 112
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 401 SQNGKNYSLKASSTKSSLSMKSH 423

RESULT 3
NRG1_CHICK STANDARD; PRT; 602 AA.
ID NRG1_CHICK

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

1 nucleic - nucleic search, using sw model

on: January 8, 2002, 17:27:15 ; Search time 158.12 Seconds
(without alignments)
1886.853 Million cell updates/sec

file: US-09-640-041-3

effect score: 348

quence: 1 atgcacacagatcacgaaaga.....acagtcattgacacacttga 348

oring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

arched: 930621 seqs, 428662619 residues

total number of hits satisfying chosen parameters: 1861242

imum DB seq length: 0

imum DB seq length: 2000000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

tabase :
N_Geneseq_1101.*
1: /SID2/gcgdata/geneseq/NA1980.DAT.*
2: /SID2/gcgdata/geneseq/NA1981.DAT.*
3: /SID2/gcgdata/geneseq/NA1982.DAT.*
4: /SID2/gcgdata/geneseq/NA1983.DAT.*
5: /SID2/gcgdata/geneseq/NA1984.DAT.*
6: /SID2/gcgdata/geneseq/NA1985.DAT.*
7: /SID2/gcgdata/geneseq/NA1986.DAT.*
8: /SID2/gcgdata/geneseq/NA1987.DAT.*
9: /SID2/gcgdata/geneseq/NA1988.DAT.*
10: /SID2/gcgdata/geneseq/NA1989.DAT.*
11: /SID2/gcgdata/geneseq/NA1990.DAT.*
12: /SID2/gcgdata/geneseq/NA1991.DAT.*
13: /SID2/gcgdata/geneseq/NA1992.DAT.*
14: /SID2/gcgdata/geneseq/NA1993.DAT.*
15: /SID2/gcgdata/geneseq/NA1994.DAT.*
16: /SID2/gcgdata/geneseq/NA1995.DAT.*
17: /SID2/gcgdata/geneseq/NA1996.DAT.*
18: /SID2/gcgdata/geneseq/NA1997.DAT.*
19: /SID2/gcgdata/geneseq/NA1998.DAT.*
20: /SID2/gcgdata/geneseq/NA1999.DAT.*
21: /SID2/gcgdata/geneseq/NA2000.DAT.*
22: /SID2/gcgdata/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Sult No.	Score	Query Match	Length DB	ID	Description
1	348	100.0	348	22	AAA91294
2	348	100.0	348	22	AAA91294
3	348	100.0	348	22	AAA91294
4	348	100.0	348	22	AAA91294
5	348	100.0	348	22	AAA91294
6	348	100.0	348	22	AAA91294
7	348	100.0	348	22	AAA91294
8	348	100.0	348	22	AAA91294
9	348	100.0	348	22	AAA91294
10	348	100.0	348	22	AAA91294
11	348	100.0	348	22	AAA91294

12	29.8	8.6	2538	20	AA06987	Mouse neurogulin r
13	29.6	8.5	5864	22	AA510458	Human stem cell fa
14	29.6	8.5	5864	22	AA41342	Human stem cell fa
15	29.6	8.5	5864	22	AA504122	Human SCF (stem ce
16	29.6	8.5	5864	22	AA504223	Human SCF (stem ce
17	29.6	8.5	5864	22	AA23900	Human SCF (stem ce
18	29.6	8.5	5864	22	AA89102	Human stem cell fa
19	29.4	8.4	462	21	AA56228	Human radiata tran
20	29.4	8.4	564	22	AA139710	Probe #8396 used t
21	29.2	8.4	352	21	AA13769	Human stem cell fa
22	29.2	8.4	353	16	AA04888	Human SCF genomic
23	29.2	8.4	1215	21	AA34773	Barley sucrose pho
24	29.2	8.4	3807	22	AA510456	Arabidopsis thalia
25	29.2	8.4	3807	22	AA41340	Human stem cell fa
26	29.2	8.4	3807	22	AA504120	Human genomic SCF
27	29.2	8.4	3807	22	AA504221	Human genomic SCF
28	29.2	8.4	3807	22	AA23898	Human genomic SCF
29	29.2	8.4	3807	22	AA89100	Human stem cell fa
30	29.2	8.4	4063	12	AAQ1540	Human Stem Cell fa
31	29.2	8.4	7253	21	AA92499	Haemophilus influe
32	29.2	8.4	7291	17	AA74176	Haemophilus adhesi
33	29.2	8.3	376	18	AA78333	Staphylococcus aur
34	29	8.3	523	21	AA38468	Arabidopsis thalia
35	29	8.3	753	21	AA35184	Arabidopsis thalia
36	29	8.3	760	21	AA51933	Arabidopsis thalia
37	29	8.3	308	22	AA31919	Human olfactory re
38	29	8.3	308	22	AA31919	Human olfactory re
39	28.8	8.3	822	20	AA48641	Human secreted pro
40	28.8	8.3	1389	20	AA17120	Human ceramidease K
41	28.8	8.3	1430	21	AA24191	Human gene express
42	28.8	8.3	7980	19	AA44243	Murine PG-Lb CDNA.
43	28.6	8.2	975	22	AA31732	Lettuce resistance
44	28.6	8.2	1260	22	AA99677	Human olfactory re
45	28.6	8.2	1260	22	AA99677	Human protein enco

ALIGNMENTS

RESULT 1
ID AAA91294 standard; DNA: 348 BP.
AC AAA91294;
XX
XX 19-JUN-2001 (first entry)
XX
XX
DE Human EGFH2 coding sequence.
XX
XX EGFH2; epidermal growth factor H2; EGFH2 protein modulated disorder;
KW cancer; peripheral neuropathy; amyotrophic lateral sclerosis; therapy;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
KW ischaemic stroke; brain injury; acute spinal cord injury; infection;
KW nervous system injury; multiple sclerosis; dementia; epilepsy;
KW peripheral nerve injury; acoustic trauma; human; ds.
XX
XX Homo sapiens.
OS
XX
XX
FH
FH Key Location/Qualifiers
FT 1..348
FT /tag= a
FT /product= "EGFH2"
CDS
XX
XX MO200114415-A2.
XX
XX 01-MAR-2001.
XX
XX 15-AUG-2000; 2000MO-US22326.
XX
XX 20-AUG-1999; 99US-0149986.
XX
XX (CHIR) CHIRON CORP.
XX

CC	-1-	DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE DOMAIN.
CC	-1-	PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR FORM (BY SIMILARITY).
CC	-1-	PTM: EXTENSIVE GLYCOSYLATION PRECEDES PROTEOLYTIC CLEAVAGE (BY SIMILARITY).
CC	-1-	SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC	-1-	SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC	-1-	SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
DR	EMBL:	U06612; AAB71812.1; ..
DR	HSSP:	Q12784; 1HRE.
DR	Interpro:	IPR000561; EGF-like.
DR	Interpro:	IPR003598; I9_c2.
DR	Interpro:	IPR003006; I9_MHC.
DR	Interpro:	IPR002154; Neuregulin.
DR	Pfam:	PF000008; EGF_1.
DR	Pfam:	PF00047; I9_1.
DR	Pfam:	PF02158; Neuregulin_1.
DR	PRINTS:	PRO1089; NEUREGULIN.
DR	SMART:	SM00181; EGF_1.
DR	SMART:	SM00408; I9c2_1.
DR	PROSITE:	PS00032; EGF_1; UNKNOWN_1.
DR	PROSITE:	PS01186; EGF_2_1.
KW	Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein; Transmembrane; Alternative splicing.	
KW	Transmembrane; Alternative splicing.	
FT	PROPEP	1 13
FT	CHAIN	14 461
FT	CHAIN	14 241
FT	DOMAIN	14 242
FT	TRANSMEM	243 265
FT	DOMAIN	266 461
FT	DOMAIN	50 119
FT	DOMAIN	165 177
FT	DOMAIN	178 222
FT	DISULFID	57 112
FT	DISULFID	182 196
FT	DISULFID	190 210
FT	DISULFID	212 221
FT	CARBOHYD	73 73
FT	CARBOHYD	120 120
FT	CARBOHYD	126 126
FT	CARBOHYD	164 164
Q	SEQUENCE	461 AA: 50890 MW; 935C9560F7148336 CRC64;

```

Query Match 20.9%; Score 130.5; DB 11; Length 461;
Best Local Similarity 25.8%; Pred.No. 1,4e-06;
Matches 33; Conservative 25; Mismatches 45; Indels 25; Gaps

QY 3 TDHEPGPSHKSFCLNGGICGVIPPISP---FCRCVENVYGCARCEEVFLPGSSIOKRS 59
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 176 TSHLYKCAEKEKFCVNGGCECFVNDLSNPSYLYCKCQGFEGACITE-NP-KMVOTQE 233
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 60 NLEPAVALAVLT-----LIIGAFYLC-----RKGFQASSVOYDINLVE 102
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 234 KASELYOKRVLTITIGIALLVG---IMCVAYCKTKKQKQKLDRLDLOSLSRSENNMV 290
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 103 TSSSAHH 110
      | |
Db 291 NIANGPHH 298

RESULT 2
Q9BD00 PRELIMINARY; PRT: 1214 AA.
Q9BD00;
DT 01-JUN-2001 (Tremblrel, 17, Created)
DT 01-JUN-2001 (Tremblrel, 17, last sequence update)
DT 01-JUN-2001 (Tremblrel, 17, last annotation update)
DE BIDRPMAL GROWTH FACTOR PRECURSOR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1].
RP SEQUENCE FROM N.A.
RA Kim J.G., Vallet J.L., Christenson R.K.;
RT "Characterization of uterine epidermal growth factor during pregnancy
RT in pigs";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF336151; AKR18830.1; -
FT CHAIN 970 .. 1022 EPIDERMAL GROWTH FACTOR.
SQ SEQUENCE 1214 AA: 133505 MW; 83D10409057458CF C6C64;

Query Match	20.68;	Score 128.5;	DB 6;	Length 1214;
Best Local Similarity	33.78;	Pred. No. 6.2e-06;		
Matches 29; Conservative	14;	Mismatches 34;	Indels 9;	Gaps 2

[illegible]

```

RESULT      3
Q9ESA9
ID Q9ESA9  PRELIMINARY;      PRT; 298 AA

```

```
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE SMDP NEUREGULIN ALPHA 2B (FRAGMENT).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
OX NCBI_TaxID=10116;
BX

RT "Structural and Functional Diversity of SMDF Neuregulin Splice
RT Variants Expressed in the Adult Rat Nervous System.";
RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
DB EMBL: AB044401; NCBI: AF0329426.1.

DR InterPro; IPR0000561; EGF-like.
DR InterPro; IPR002154; Neuregulin.
DR InterPro; IPR002114; Pts_HPR_ser.
DR Pfam; PF000068; EGF_1.
DR Pfam; PF02155; Neuregulin; 1.
DR PRINTS; PRO1089; NEUREGULIN.
DR SMART; SM00181; EGF_1.
DR SMART; SM00001; EGF_like; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00585; Pts_HPR_ser; UNKNOWN_1.
FT NON_TER 1 1
FT NON_TER 298 298
SQ SEQUENCE 298 AA; 32851 MW; BD76F014C2B33026 CRR64;

Query match	19.8%	Score 123.57	DB 11	Length 298
Best Local Similarity	25.0%	Pred. NO. 5.3e-06		
Matches	32	Conservative	25	Mismatches 46; Indels 25; Gaps 6;

Db 46 TSHLICKAEKKITCVANGECFTYKDLNSPRYLCKOPOGTGRCTE -NVP -MKVOTOE 1030
QY 60 NLEFAVALAVLT-----LIIGAFYLC-----RKGHFORASSVOYDINLVE 1021
Db 104 KAEELYOKRVLITIGICALLVG---IMCVAYCKTKKROKHLDRLOSLRESENVL 160

Novel epidermal growth factor H2 (EGFH2) polypeptides and polynucleotides encoding them useful for treating EGFH2 protein-modulated disorders such as cancers, Alzheimer's disease, ischemic stroke and epilepsy

Claim 10; Fig 1; 48pp: English.

This sequence is the human epidermal growth factor H2 (EGFH2) protein of the invention. The EGFH2 DNA and protein sequences are useful for modulating the amount of EGFH2 protein in a subject, or treating an EGFH2 protein modulated disorder. The DNA or its antisense construct, ribozyme or a retroviral vector comprising a promoter are also useful for treating breast, prostate, pancreatic, oral or ovarian cancers, peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's disease, ischemic stroke, brain injury, acute spinal cord injury, nervous system injury, multiple sclerosis, infection, dementia, epilepsy, peripheral nerve injury, acoustic trauma and tissue wound. The DNA and protein are also useful for providing trophic support for cells in a patient by implanting cells expressing them encapsulated in a semipermeable membrane. EGFH2 polypeptides can be used to screen peptide libraries. Antibodies which bind to EGFH2 and/or variant polypeptides can be used in diagnosing and determining the prognosis of cancer, tumour progression, hyperproliferative cell growth or accompanying biological and physical manifestations.

Sequence 115 AA:

Query Match 100.0%; Score 625; DB 22; Length 115;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MPTDHEPCGSHKSFCLNGGLCYVPTIPSPRCVNTGARCEVFLPGSSIOGTSN 60
|||||
1 mptdhepcgshksfclngglcyvptipsprcvntgarceevflpgssioqtskn 60
61 LFEAFVALAVLTLLIGAFYFLCRKGHPORASSVOYDINLVETSSSAHSHSQH 115
|||faafvalavltlligafyflcrkghfporassvoydinlvetsstahshsqh 115

RESULT 2
ID AAY97698
AAY97698 standard; Protein; 115 AA.

AC AAY97698;

DT 19-JUN-2001 (first entry)

DE Mouse EGFH2 protein sequence.

EGFH2: epidermal growth factor H2; EGFH2 protein modulated disorder: cancer; peripheral neuropathy; amyotrophic lateral sclerosis; therapy; Alzheimer's disease; Parkinson's disease; Huntington's disease; ischemic stroke; brain injury; acute spinal cord injury; infection; nervous system injury; multiple sclerosis; dementia; epilepsy; peripheral nerve injury; acoustic trauma; mouse.

OS mus sp.

PN WO200114415-A2.

PD 01-MAR-2001.

PF 15-AUG-2000; 2000MO-US22326.

PR 20-AUG-1999; 990S-0149986.

PA (CHIR) CHIRON CORP.

PI Kavanagh WM, Cen H, Lee P;

DR WPI: 2001-211304/21.
DR N-PSDB; AAA91295.

Novel epidermal growth factor H2 (EGFH2) polypeptides and polynucleotides encoding them useful for treating EGFH2 protein-modulated disorders such as cancers, Alzheimer's disease, ischemic stroke and epilepsy

Claim 10; Fig 1; 48pp: English.

This sequence is the mouse epidermal growth factor H2 (EGFH2) protein of the invention. The EGFH2 DNA and protein sequences are useful for modulating the amount of EGFH2 protein in a subject, or treating an EGFH2 protein modulated disorder. The DNA or its antisense construct, ribozyme or a retroviral vector comprising a promoter are also useful for treating breast, prostate, pancreatic, oral or ovarian cancers, peripheral neuropathy, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, Huntington's disease, ischemic stroke, brain injury, acute spinal cord injury, nervous system injury, multiple sclerosis, infection, dementia, epilepsy, peripheral nerve injury, acoustic trauma and tissue wound. The DNA and protein are also useful for providing trophic support for cells in a patient by implanting cells expressing them encapsulated in a semipermeable membrane. EGFH2 polypeptides can be used to screen peptide libraries. Antibodies which bind to EGFH2 and/or variant polypeptides can be used in diagnosing and determining the prognosis of cancer, tumour progression, hyperproliferative cell growth or accompanying biological and physical manifestations.

Sequence 115 AA:

Query Match 78.2%; Score 489; DB 22; Length 115;
Best Local Similarity 75.7%; Pred. No. 1e-38;
Matches 87; Conservative 12; Mismatches 16; Indels 0; Gaps 0;

1 MPTDHEPCGSHKSFCLNGGLCYVPTIPSPRCVNTGARCEVFLPGSSIOGTSN 60
|||||
1 mptdhepcgshksfclngglcyvptipsprcvntgarceevflpgssioqtskn 60
61 LFEAFVALAVLTLLIGAFYFLCRKGHPORASSVOYDINLVETSSSAHSHSQH 115
|||faafvalavltlligafyflcrkghfporassvoydinlvetsstahshsqh 115

RESULT 3
ID AAW97617
AAW97617 standard; Protein; 713 AA.

AC AAW97617;

DT 10-MAY-1999 (first entry)

DE Mouse neuregulin related ligand NRG3.

Neuregulin related ligand; NRG3; mouse; ErbB4 receptor; signal transduction; nervous system disorder; neurodegeneration; neuropathy; therapy; diagnosis.

OS Mus sp.

PN Location/Qualifiers

FT 1..362
FT Domain
FT /note= "extracellular domain, specifically claimed in Claim 5(a)"

FT 66..91
FT Region
FT /note= "hydrophobic region"

FT 105..286
FT Region
FT /note= "mucin-like Ser/Thr-rich region, contains sites for O-linked glycosylation"

FT 287..334
FT Domain
FT /note= "EGF-like domain"

FT 363..385
FT Domain

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

nucleic - nucleic search, using sw model

January 8, 2002, 15:14:05 ; Search time 1410.3 seconds
(without alignments)
4070.778 Million cell updates/sec

US-09-640-041-3

fect score: 348
1 atgcacacagatcacgaaga.....acagtcacgacacacatga 348

quence:

ring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

atched: 1472140 segs, 8248589755 residues

tal number of hits satisfying chosen parameters: 2944280

imum DB seq length: 0
imum DB seq length: 2000000000

st-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

tabase :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_hcg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_pro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_om:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_pat:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_sy:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_hcgo_hum:*
- 31: em_hcgo_inv:*
- 32: em_hcgo_rod:*
- 33: em_hcg_hum:*
- 34: em_hcg_inv:*
- 35: em_hcg_rod:*
- 36: em_hcg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	348	100.0	348	6	AX088172	AX088172 Sequence
2	252	72.4	348	6	AX088170	AX088170 Sequence
3	252	72.4	750	10	AF083067	AF083067 Mus muscu
4	157.8	45.3	166352	2	AC026180	AC026180 Homo sapi
5	157.8	45.3	171000	2	AC087456	AC087456 Homo sapi
6	157.8	45.3	189076	2	AC068838	AC068838 Homo sapi
7	101.2	29.1	154323	2	AC026033	AC026033 Homo sapi
8	101.2	29.1	161262	2	AC087464	AC087464 Homo sapi
9	101.2	29.1	166352	2	AC026180	AC026180 Homo sapi
10	90	25.9	180795	2	AC090168	AC090168 Homo sapi
11	48.6	14.0	188553	2	CNS01RH3	AL160314 Homo sapi
12	37.2	10.7	204779	4	BOVBUTB1	AL156314 Homo sapi
13	35.4	10.2	2691	4	AC073727	M55551 Bovine buty
14	35.4	10.2	170139	9	AL136220	AC073727 Mus muscu
15	35	10.1	150577	2	AC073721	AL136220 Human DNA
16	35	10.1	190577	2	AC073721	AC073721 Mus muscu
17	35	10.1	199684	9	AC005036	AC005036 Homo sapi
18	35	10.1	241714	2	AF108420	AC079564 Mus muscu
19	34.8	10.0	46626	5	AF108420	AF108420 Fugu rubr
20	34.6	9.9	154242	2	AC073373	AC073373 Mus muscu
21	34.6	9.9	163542	9	AL158206	AL158206 Human DNA
22	34.2	9.8	128294	9	AB023558	AB023558 Human ade
23	34.2	9.8	163117	8	AC018733	AC0733248 Homo sapi
24	34.2	9.8	163117	8	SPU73044	AC018733 Homo sapi
25	33.8	9.7	2855	8	SPBC1861	U73044 Schizosach
26	33.8	9.7	23851	8	AC044819	AC044819 Homo sapi
27	33.6	9.7	151798	2	AC069199	AC069199 Homo sapi
28	33.6	9.7	159991	2	AL162734	AL162734 Homo sapi
29	33.6	9.6	37968	8	CAC20C1	AL033391 C. albican
30	33.4	9.6	75382	2	AC023242	AC023242 Homo sapi
31	33.4	9.5	224863	2	AC093165	AC093165 Homo sapi
32	33.2	9.5	177767	2	AC007342	AC007342 Homo sapi
33	33	9.5	179202	2	AC025935	AC025935 Homo sapi
34	33	9.5	188863	9	AC007345	AC007345 Homo sapi
35	33	9.5	190963	2	AC023597	AC023597 Homo sapi
36	32.8	9.4	17867	2	AC017551	AC017551 Homo sapi
37	32.8	9.4	36654	2	AC012365	AC012365 Homo sapi
38	32.8	9.4	126908	2	AP002432	AP002432 Homo sapi
39	32.8	9.4	139515	2	AP002432	AP002432 Homo sapi
40	32.8	9.4	159804	2	AP000793	AP000793 Homo sapi
41	32.8	9.4	169152	9	AC005714	AC005714 Drosophila
42	32.8	9.4	177740	2	AC021901	AC021901 Homo sapi
43	32.8	9.4	181893	3	AC008348	AC008348 Drosophila
44	32.8	9.4	191466	3	AC008348	AC008348 Drosophila
45	32.8	9.4	191466	3	AC008348	AC008348 Drosophila

ALIGNMENTS

RESULT 1

AX088172 AX088172 348 bp DNA
LOCUS Sequence 3 from Patent WO0114415.
DEFINITION AX088172
ACCESSION AX088172
VERSION AX088172.1 GI:13397084
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

PAT 17-MAR-2001

REFERENCE 1 (bases 1 to 348)
AUTHORS Kavanagh, W.M., Gen, H. and Lee, P.
TITLE Egf2 genes and gene products
JOURNAL Patent: WO 0114415-A 3 01-MAR-2001;
CHIRON CORPORATION (US)

FEATURES
source
1..348
/organism="Homo sapiens"
/db_xref="taxon:9606"

266 SEHFKPCROKDLAYCLINDECFYETLTGSHKCRKCGYGVRCDO-FLPKTDSILSDP 344
52 -----GSSIOTKNLE-----AFVALAVLTLLIGAFYFLCRKGHFORASSVO----- 95
345 TDHLIEFMESVEDYOROVLSICITIFGIYVGMFCALFYKSKR-----QAKOIEHLKE 400
Y 96 -----YDINLVETSSSAHSH 112
b 401 SONGKNYSLKASSTKSESILMSH 423

RESULT 2
US-09-126-121-2
Sequence 2, Application US/09126121
Patent No. 6252051

GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
TITLE OF INVENTION: Ligands and Uses Therefor
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/126,121
FILING DATE: 30-Jul-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Deirdre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 713 amino acids
TYPE: Amino Acid
TOPOLOGY: linear
FEATURE:
NAME/KEY: Mouse NRG3 (mNRG3)/amino acid seq.
LOCATION: 1-713
IDENTIFICATION METHOD:
OTHER INFORMATION:
US-09-126-121-2

Query Match 22.7%; Score 142; DB 4; Length 713;
Best Local Similarity 26.6%; Pred. No. 6,1e-08;
Matches 38; Conservative 20; Mismatches 47; Indels 38; Gaps 6;
OY 3 TDHEPCGSPKSHKSFCLNGGCGYVPTIPSP---FCRCVENYTGARCEVPF----- 51
Db 286 SEHFKPCROKDLAYCLINDECFYETLTGSHKCRKCGYGVRCDO-FLPKTDSILSDP 344
OY 52 -----GSSIOTKNLE-----AFVALAVLTLLIGAFYFLCRKGHFORASSVO----- 95
Db 345 TDHLIEFMESVEDYOROVLSICITIFGIYVGMFCALFYKSKR-----QAKOIEHLKE 400
OY 96 -----YDINLVETSSSAHSH 112
Db 401 SONGKNYSLKASSTKSESILMSH 423

RESULT 3
US-08-168-091A-2
Sequence 2, Application US/08168091A
Patent No. 5665862
GENERAL INFORMATION:
APPLICANT: Fischbach, Gerald.
APPLICANT: Falls, Douglas R.
APPLICANT: Rosen, Kenneth M.
APPLICANT: Cortis, Gabriel.
TITLE OF INVENTION: Neurotrophic Factor
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE AND COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/168,091A
FILING DATE: 15-DEC-1993
CLASSIFICATION: 435.
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/953,742
FILING DATE: 29-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: HMI-002CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 602 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-168-091A-2

Query Match 22.4%; Score 140; DB 1; Length 602;
Best Local Similarity 27.1%; Pred. No. 8.3e-08;
Matches 35; Conservative 21; Mismatches 51; Indels 22; Gaps 4;
OY 3 TDHEPCGSPKSHKSFCLNGGCGYVPTIPSP---FCRCVENYTGARCEVPF----- 48
Db 135 TSHLTCKDIOKAFYVNGGCGYVWKPDPNPRYLRCRPNFTGRCRNYWASFYKHLGI 194
OY 49 -FLPGSSIOTKNLEFEAFVALAVLTLLIGAFYFLC-----RKGHFORASSVOYDINLV 101
Db 195 EFMEAEELVQKRVLTITGICITALLVGMCVAV-CKTKRQKKLHRLQSLRSENNV 253
OY 102 ETSTSTSAH 110
Db 254 MNMANGPH 262

RESULT 4
US-08-428-926-5
Sequence 5, Application US/08428926
Patent No. 5667780
GENERAL INFORMATION:
APPLICANT: Ho, Wei-Hsien
APPLICANT: Osheroff, Phyllis L.
TITLE OF INVENTION: SENSOR AND MOTOR NEURON DERIVED FACTOR (SMDF)
NUMBER OF SEQUENCES: 5